REMARKS/ARGUMENTS

Reconsideration and allowance of this application are respectfully requested. Currently, claims 1-11, 13-15, 17-45, 47 and 49 are pending in this application.

Rejections Under 35 U.S.C. §102 and §103:

Claims 1-8, 20-24, 26-27 and 34-36 were rejected under 35 U.S.C. §102(e) as allegedly being anticipated by Gilai et al (U.S. '736, hereinafter "Gilai"). Applicant respectfully traverses this rejection.

For a reference to anticipate a claim, each element must be found, either expressly or under principles of inherency, in the reference. Applicant submits that Gilai fails to disclose each element of the claimed invention. For example, Gilai fails to disclose or even suggest the following elements of independent claim 1:

> "control means operable:
> (a) to control the speech recognition means to identify, by reference to recognition information for the first set of words, as many words of the first set as meet a predetermined criterion of similarity to first received voice signals;
> (b) upon such identification, to compile a list of all words of the second set which are connected with entries connected also with the identified word(s) of the first set; and
> (c) to control the speech recognition means as to identify, by reference to recognition information for the second set of words, at least one word of the list which resembles second received voice signals."

Gilai discloses a database access system which receives a speech signal as input to a speech recognizer (Figure 11, col. 6, line 2 and col. 18, lines 35-45) and outputs phonemic representations of the input speech, together with associated probabilities of

each phoneme having occurred (col. 18, ll. 44-63). As shown in Figure 11, these phoneme outputs are input to the database access system 20, where the spellguess unit 30 receives them as input (Figure 1). The spellguess unit 30 selects one or more strings of symbols to which the phoneme input might relate, each associated with a probability (col. 6, ll. 5-8). These selected strings and their probabilities are then output to a phonetic translator 40, which may translate each of them into phonetic form (Figure 1 and col. 6, ll. 17-19). Alternatively, the phonetic translator 40 may be omitted, in which case the output strings are input to a dictionary look-up unit 50 (col. 6, ll. 42-45).

For each received string, the dictionary look-up unit 50 conducts a similarity search within a dictionary 60 (col. 6, ll. 47-49). For each received string, the dictionary look-up unit 50 thus looks up similar strings, and outputs for each string: 1) a possibly empty list of dictionary entries whose similarity to the string is better than or equal to a predetermined similarity threshold, and 2) a similarity score representing the degree of similarity between that entry and the corresponding string (col. 7, ll. 11-17).

The list of dictionary entries obtained from the dictionary look-up unit 50 are then passed to a database look-up unit 90 (Figure 1). The database unit look-up unit 90 identifies, for each dictionary entry within the list, database entries which include the dictionary entry (col. 7, ll. 33-35). The database look-up unit 90 operates using a concordance 72, which is a list of entries including a single entry for each occurrence of each dictionary word in the database, and a pointer to the database entry for that particular occurrence (col. 9, ll. 3-5). Each database entry is pointed to by all

concordance entries corresponding to that database entry (col. 9, ll. 10-11 and Figure 12). For any particular word supplied in the list of dictionary entries from the dictionary entry look-up unit 50, the database look-up unit 90 may therefore follow the pointers stored within the concordance so as to identify corresponding database entries thereto.

As the database look-up unit 90 operates, it maintains a similarity vector 92 which has a number of vector components corresponding to the number of database entries (col. 7, ll. 46-50). Each similarity vector component stores the similarity score of the corresponding database entry, the similarity score being provided for each word by the dictionary look-up unit 50 as discussed above. The database look-up unit 90 accumulates the similarity scores as each word is looked up (col. 7, ll. 50-55).

The similarity vector component scores are used by a candidate sorter 100 to extract a predetermined number of the highest scoring similarity vector components for storage in a best candidates box 94 (col. 7, ll. 57-61). A conversation manager 80 then receives the output of the database look-up unit 90 and the candidate sorter 100, and produces an output for the user depending on the number of database entries identified. The available outputs are described at col. 8, ll. 5-33. In particular, col. 8, ll. 10-20 discloses that if each database entry includes several fields, such as street name, house number, locality, etc., and a single database entry has not been identified, the conversation manager 80 may decide to prompt the user to enter information identifying another word within the desired database entry. In this case, the database look-up unit 90 operates for each database entry component so that after a further word has been elicited

from a user, the database look-up unit 90 updates the similarity vector 92 such that the similarity vector accumulates a similarity score **for each database entry**.

In summary, Gilai describes a database access system which uses a speech recognizer to receive speech input of words spoken by a user and then looks-up database entries depending upon the recognized word. A dictionary look-up unit looks-up similar words to the input word, and produces a similarity score for each identified similar word. Each identified similar word is then used as input to a database look-up unit, which uses a concordance between the dictionary and the database entries to identify database entries that contain the identified dictionary words For each dictionary word for which a database entry is looked up, the similarity score is stored in a similarity vector component for that database entry. As each identified dictionary word is looked up to find the corresponding database entry, the similarity vector components of the identified database entries are updated and accumulated if necessary. Therefore, for example, if the database look-up unit identifies a first database entry when looking up a particular identified dictionary word, the similarity score for that identified dictionary word as determined by the dictionary look-up unit is stored in the similarity vector component for that database entry. If, when processing the next identified dictionary word, the same database entry is identified, the similarity score for the second identified dictionary word is accumulated with the previous similarity score in the particular database entry's similarity vector component.

Database entries therefore accumulate similarity scores in their respective similarity vector components as each word identified by the dictionary look-up unit is processed by the database look-up unit 90. A candidate sorter 100 may identify those database entries with the highest scoring similarity vector components as possible outputs by the conversation manager 80. If a particular database entry has a sufficiently high similarity score then it may be output to the user. Alternatively, if there is no single database entry which stands out as a possible result, then the user is asked to elicit another input word, and the process recommences. The similarity vector components **for each** database entry are updated during processing of the second user response. If a second word is elicited from the user, every word in the dictionary is therefore available to be determined as a similar word thereto, and each database entry may correspondingly have its similarity vector component score updated in response thereto.

Gilai therefore fails to disclose, *inter alia*, control means operable: a) to control the speech recognition means to identify words of a first set which meet a predetermined criterion of similarity to a first voice signal, b) to compile a list of words of a second set which are connected with database entries connected also with the identified words of the first set, and c) control the speech recognition means to identify at least one word of the list which resembles second received voice signals as required by claim 1. Claim 1 thus requires first received voice signals being recognized in accordance with the first set of words, and second received voice signals being recognized in accordance with a list of particular words of a second set of words which are connected with entries connected

also with the identified words of the first set. Thus, for each "round" of speech recognition, a different vocabulary is used by the speech recognizer. In contrast, Gilai discloses no such control of the vocabulary of the speech recognizer. At each recognition operation, the full speech recognizer vocabulary is available. Gilai thus fails to disclose each element required by claim 1.

Independent claim 20 requires, *inter alia*, speech recognition means operable upon receipt of a first voice signal to identify as many words of a first set as meet a predetermined recognition criterion, means to generate a list of all words of a second set which are related to an entry to which the identified word(s) of the first set is also related, and speech recognition means operable upon receipt of a second voice signal to identify at least one word of the list. Independent claim 21 requires, *inter alia*, recognition means operable upon receipt of a first input pattern signal to identify as many patterns of a first set as meet a predetermined recognition criterion, means to generate a list of all patterns of a second set which are related to an entry to which an identified pattern of the first set is also related, and recognition means operable upon receipt of a second input pattern signal to identify at least one pattern of the list. Gilai fails to disclose these limitations of claims 20 and 21 for reasons similar to those discussed above with respect to claim 1.

Independent claim 22 requires, *inter alia*, means for identifying a received signal as corresponding to as many of a first set of signals as meet a predetermined criterion, control means operable to compile a list of all words of a second set which are connected with entries connected also with the identified signal of the first set, and speech

recognition means operable to identify by reference to recognition information for the second set of words at least one word of the list which resembles received voice signals. Gilai fails to disclose or even suggest the above claimed features. A speech recognition means which has had its vocabulary dynamically reduced depending upon previously received signals is used to identify further received voice signals through the above claimed features. Nowhere in Gilai is a dynamically reduced vocabulary of a speech recognizer disclosed.

Independent claim 26 requires, *inter alia*, compiling a list of those words connected with entries connected also with identified words, and identifying at least one of the words of the list as present in the received voice signals. Independent claim 27 requires similar features. Gilai fails to disclose restricting identification of received voice signals to words contained within a list which has been compiled depending upon the identification of previously received voice signals. Gilai thus fails to disclose each limitation required by claims 26 and 27.

Independent claim 34 and claims 36-37 which depend therefrom require, *inter alia*, compiling a list of second sub-items mutually interrelated with identified first sub-item(s), and performing speech recognition of second speech input with respect to the compiled list to identify at least one potentially corresponding second sub-item from the list. Within Gilai, after a first speech input has been received, a list of database entries with the highest scoring similarity vector components is compiled by the candidate sorter 100 (col. 7, ll. 59-62), but if no single database entry stands out as a result then the user

may be prompted for a further speech input. However, as specified at col. 15, ll. 31-34, after each operation (in this case the elicitation of further input from a user) the database look-up unit 90 updates the similarity vector such that the similarity vector accumulates a similarity score **for each** database entry. The candidate sorter 100 compiles a further list based upon the newly accumulated similarity vector component similarity scores, irrespective of whether a database entry was contained within the previous list generated by the candidate sorter 100. Therefore, Gilai does not disclose restricting identification of database items in response to the second speech input to those which are identified in response to the first speech input. In contrast to claim 34, each database entry of Gilai is available to be identified by the candidate sorter 100.

Accordingly, Applicant submits that claims 1-8, 20-24, 26-27 and 34-36 are not anticipated by Gilai and respectfully requests that the rejection of these claims under 35 U.S.C. §102 be withdrawn.

Claims 12, 14, 17, 38, 40, 42-43 and 46-48 were rejected under 35 U.S.C. §102(e) as allegedly being anticipated by Lennig et al (U.S. '488, hereinafter "Lennig"). Applicant respectfully traverses this rejection with respect to still pending claims 14, 17, 38, 40, 42-43 and 47.

Lennig describes an automated directory assistance apparatus having a voice processing unit. The voice processing unit is provided with several lexicons each comprising a group of lexemes having common characteristics (abstract and col. 5, ll. 53-

58). A different lexicon is used by the voice processing unit depending upon the state of progress of the call, and in particular on the prompt which it has just issued to the caller (col. 5, ll. 58-61). In one embodiment, an a priori probability function is used in a directory assistance query to weight each lexeme in a selected lexicon based on the locality of the calling number (col. 5, ll. 48-52 and col. 6, ll. 26-31). The speech recognition process then involves an acoustic determination based upon the acoustics of the user response, together with the assigned a priori probability for each lexeme (col. 7, ll. 21-24).

Determination of the probability function is described in col. 7, line 55 to col. 9, line 62. This probability function does not tend to zero, but instead tends to 0.01, above a distance value of 167 (col. 9, ll. 55-59). This means that each lexeme in the lexicon which is being used by the speech recognizer is available for recognition by the speech recognizer, even if it has been assigned an a priori probability of only 0.01. That is, there is no restriction as to which lexemes within a particular lexicon may be recognized by the speech recognizer by the use of the a priori probability function.

In a second embodiment described from col. 10, line 62 to col. 11, line 11, instead of computing a probability index for all possible localities served by the directory assistance apparatus, the locality lexemes within the locality lexicon may be grouped into predetermined subsets according to call identifiers. When a directory assistance call is received, a particular one of the subset of localities is selected in dependence upon the call identifier and the speech recognition process is limited thereto. Each subset is pre-

selected to include those localities which give greatest recognition accuracy during actual service.

Independent claim 14 requires, *inter alia*, "the control means is operable to generate recognition data for each word of the subset." Lennig fails to disclose, or even suggest, this claimed feature. In particular, Lennig discloses lexicon subsets which are predetermined (col. 10, ll. 65-67). Lennig thus fails to disclose any control means which is operable to generate recognition data for each word of the subset as required by claim 14.

Independent claim 17 requires "wherein means responsive to receipt via the telephone line connection of signals indicating the origin of a telephone call to access stored information identifying one of the sets of recognition data and to supply this set to the recogniser; wherein the stored sets correspond to different languages or regional accents." While Lennig may describe identifying sets of recognition data used by a speech recognizer in response to the origin of a telephone call, within Lennig those sets (lexicon subsets within Lennig) are subsets of locality data (col. 11, line 4), and not sets corresponding to different languages or regional accents as required by claim 17. While Lennig discloses the use of bilingual prompts (col. 6, ll. 1-4), no selection of such prompts is made dependent upon the origin of a telephone call. Lennig therefore fails to disclose each limitation recited in claim 17.

Independent claim 38 and claim 40 which depends therefrom require "means responsive to receipt via the telephone line connection of signals indicating the

destination of a telephone call to access stored information identifying a subset of the set of utterances and to restrict the recogniser operation to that subset." Similarly, independent claim 42 requires, *inter alia*, "means responsive to receipt via the telephone line connection of signals indicating the destination of a telephone call to access stored information identifying one of the sets of recognition data and to supply this set to the recogniser." While col. 10, line 62 to col. 11, line 11 of Lennig may disclose computing a probability index for all destination localities served by a directory assistance apparatus, Lennig fails to disclose or even suggest the identification of a subset of utterances or the selection of sets of recognition data depending on a destination of a telephone call. Lennig merely discloses an a priori probability using the calling number to determine a probability index which will be used to weight the speech recognition result.

Independent claim 47 requires, *inter alia*, determining or verifying an identity of a speaker of received spoken words by reference to recognition data corresponding to a set of possible speakers. Lennig fails to disclose this claimed feature. Lennig also fails to disclose accessing stored information identifying a subset of the set of speakers and restricting recognition operation to that the subset in response to the receipt of signals indicating the origin of the spoken words. Lennig simply does not disclose such speaker identification or verification steps.

Accordingly, Applicant submits that claims 14, 17, 38, 40, 42-43 and 47 are not anticipated by Lennig and respectfully requests that the rejection of these claims under 35 U.S.C. §102(e) be withdrawn.

Claims 18-19 and 44-45 were rejected under 35 U.S.C. §103 as allegedly being unpatentable over Lennig in view of well known prior art. Applicant respectfully traverses these rejections.

In order to establish a prima facie case of obviousness, all of the claimed limitations must be taught or suggested by the prior art. Claim 18 and claim 19 which depends therefrom require, *inter alia*, "means responsive to receipt via the telephone line connection of signals indicating the origin of a telephone call to access stored information identifying one of the sets of recognition data and to supply this set to the recogniser; wherein at least two of the sets correspond to the characteristics of different types of telephone apparatus." While it is generally known in the art to allow access to data and services using different types of telephone apparatuses, claim 18 is concerned with identifying recognition data sets depending upon the characteristics of the different types of telephone apparatuses. This feature is not taught or suggested by Lennig and the "well known prior art".

Claims 44 and 45 depend from claim 42. As discussed above with respect to claim 42, Lennig fails to disclose or even suggest "means responsive to receipt via telephone line connection of signals indicating the destination of a telephone call...." The "well known prior art" fails to remedy this deficiency of Lennig.

Accordingly, Applicant submits that claims 18-19 and 44-45 are not "obvious" over Lennig in view of well known prior art and thus respectfully requests that the rejection of these claims under 35 U.S.C. §103 be withdrawn.

Claims 9-11, 25, 28-33 and 37 have been rejected under 35 U.S.C. §103 as allegedly being unpatentable over Gilai in view of Lennig. Applicant traverses this rejection.

Claims 9-11 depend at least indirectly from independent claim 1 and claim 25 depends from independent claim 22. As discussed above, Gilai fails to disclose or even suggest (a), (b) and (c) recited in claim 1. Also as discussed above, Gilai fails to disclose or even suggest means for identifying a received signal as corresponding to as many of a first set of signals as meet a predetermined criterion, control means operable to compile a list of all words of a second set which are connected with entries connected also with the identified signal of the first set, and speech recognition means operable to identify, by reference to recognition information for the second set of words, at least one word of the list which resembles received voice signals as required by claim 22. Lennig fails to remedy these deficiencies of Gilai. Claims 9-11 and 25 are thus not "obvious" over Gilai in view of Lennig.

Claims 28-32 require, *inter alia*, compiling a list of words which is a subset of a set of stored words to be recognized as a function of a second signal, and applying to a stored speech signal a speech recognition process so as to identify, by reference to the list

at least one word of the subset. The combination of Lennig and Gilai fails to teach or suggest these claim limitations.

Independent claim 33 requires, *inter alia*, receiving and storing a speech signal, performing a recognition operation, and in the event of the recognition operation fails to meet a predetermined criterion of reliability, retrieving the stored speech signal and performing a recognition operation thereon. Claim 33 thus essentially requires a time delay of the recognition of the retrieved and stored speech signal until some other operation has been performed. The combination of Gilai and Lennig fails to teach or suggest these claimed features.

Claim 37 depends from independent claim 34 and thus requires compiling a list of second sub-items mutually interrelated with identified first sub-item(s), and performing speech recognition of second speech input with respect to the compiled list to identify at least one potentially corresponding second sub-item from the list. Lennig fails to remedy the above described deficiency of Gilai with respect to these claim limitations and thus claim 37 is not "obvious" over Gilai in view of Lennig.

Accordingly, Applicant submits that claims 9-11, 25, 28-33 and 37 are not "obvious" over Gilai in view of Lennig and thus respectfully requests that the rejection of these claims under 35 U.S.C. §103 be withdrawn.

Claims 13 and 39 were rejected under 35 U.S.C. §103 as allegedly being unpatentable over Lennig in view of Gilai. Applicant respectfully traverses this rejection.

Independent claim 13 relates to marking in a recognition data store those items of data therein which correspond to words which are or which are not in the subset, and a recognition means ignoring words so marked, or conversely not marked. In contrast, the probabilistic spoken phoneme recognition unit 770 which performs the speech recognition in the database accessing system of Gilai is completely separate from the processing performed by the database accessing unit 20 (see Fig. 11). Contrary to the allegations of the Patent Office, the database look-up unit 90 has no interaction with the probabilistic spoken phoneme recognition unit 770. The database look-up unit 90 therefore does not perform marking in a recognition data store those items which correspond to words which are or which are not in the subset as claimed. Lennig also fails to teach or suggest this feature.

Claim 39 depends from independent claim 38. As discussed above with respect to claim 38, Lennig fails to disclose means responsive to receipt via a telephone line connection of signals indicating a destination of a telephone call to access stored information identifying a subset of a set of utterances and to restrict recognizer operation to that subset. Gilai fails to remedy this deficiency of Lennig. Accordingly, even if Gilai and Lennig were combined as proposed by the Office Action, the combination would not have taught or suggested all of the claimed limitations.

Accordingly, Applicant respectfully submits that claims 13 and 39 are not "obvious" over Lennig in view of Gilai and respectfully requests that the rejection of these claims under 35 U.S.C. §103 be withdrawn.

Claims 15, 41, 47 and 49 were rejected under 35 U.S.C. §103 as allegedly being unpatentable over Borcherding (EPO '688) in view of Lennig. Applicant respectfully traverses this rejection.

Claims 15, 41 and 49 relate to using the origin of a telephone call or spoken words to identify a subset of a set of speakers, and to restrict a speech recognizer which determines or verifies the identity of a speaker to the identified subsets. The Office Action apparently alleges that Borcherding discloses a telephone line connection and a speech recognizer for determining or verifying the identity of the speaker via the telephone line connection, but does not disclose identifying a subset of the set of speakers, depending on the origin or destination of a telephone call or the spoken words. The Office Action then apparently alleges that Lennig rectifies this deficiency of Borcherding either by the use of the disclosed a priori probability function or lexicon subset of localities. However, neither the a priori probability function nor the lexicon subset of localities relates to using the origin of a telephone call to identify a subset of speakers and then subsequently restricting operation of a speech recognizer to determine or verify the identity of a spoken word to that subset. Accordingly, even if Borcherding and Lennig were combined as proposed by the Office Action, the combination would not have taught or suggested all of the claimed limitations.

Claim 47 requires, *inter alia*, "responsive to receipt of signals indicating the origin of these spoken words, accessing stored information identifying a subset of the set of speakers and restricting the recognizing operation to that subset." As discussed above,

Lennig fails to disclose or suggest this claimed feature. Borcherding also fails to disclose or suggest this feature and thus fails to remedy the deficiencies of Lennig.

Accordingly, Applicant submits that claims 15, 41, 47 and 49 are not "obvious" over Borcherding in view of Lennig and therefore respectfully requests that the rejection of these claims under 35 U.S.C. §103 be withdrawn.

Conclusion:

Applicant believes that this entire application is in condition for allowance and respectfully requests a notice to this effect. If the Examiner has any questions or believes that an interview would further prosecution of this application, the Examiner is invited to telephone the undersigned.

Respectfully submitted,

NIXON & VANDERHYE P.C.

By: ______________________
Raymond Y. Mah
Reg. No. 41,426

RYM:sl
1100 North Glebe Road, 8th Floor
Arlington, VA 22201-4714
Telephone: (703) 816-4000
Facsimile: (703) 816-4100